10 November 2016

Introduction to the Research Data Alliance for the Transportation Community

SPEAKER: Lynn Yarmey, Research Data Alliance

The Research Data Alliance

Lynn Yarmey
RDA/US Director of Community Development

yarmel@rpi.edu

@ LynnYarmey @RESDATALL @RDA_US

The Research Data Alliance

Who am I?

Edited photo, original courtesy of Flickr user Albert at http://bit.ly/1UfNoQL

What is RDA?

RDA is building the social and technical bridges that enable open sharing of data

What is RDA?

RDA *is building* the social and technical bridges that enable open sharing of data

Vision
Researchers and innovators openly share data across technologies, disciplines, and countries to address the grand challenges of society

Mission
RDA builds the **social and technical bridges** that **enable open sharing** of data

THE RESEARCH DATA ALLIANCE

www.rd-alliance.org

building the social and technical bridges that enable open sharing of data

15 FLAGSHIP OUTPUTS

of which 4 ICT Technical Specifications

75 ADOPTION CASES

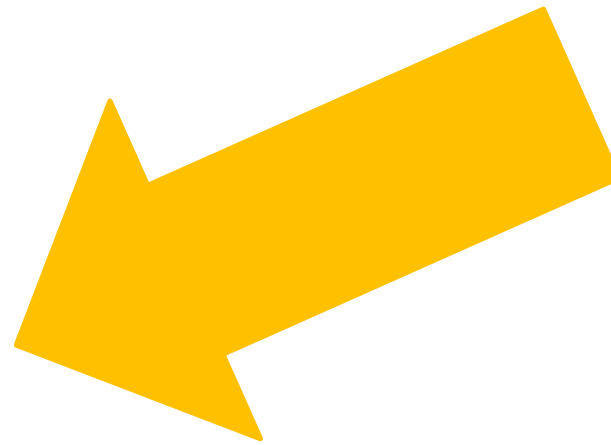across multiple disciplines, organisations & countries

72 GROUPS WORKING ON GLOBAL DATA INTEROPERABILITY CHALLENGES

of which 27 WORKING GROUPS & 45 INTEREST GROUPS

4,527 INDIVIDUAL MEMBERS FROM 115 COUNTRIES

66% Academia & Research
16% Public Administration
11% Enterprise & Industry

46 ORGANISATIONAL MEMBERS

RDA
RESEARCH DATA ALLIANCE

What does RDA do?

RDA Recommendations

Data Foundation & Terminology: a model for data in the registered domain.

PID Information Types: a common protocol for providers and users of persistent ID services worldwide.

Data Citation: defining mechanisms to reliably cite dynamic data

Practical Policy: defining best practices of how to deal with data automatically and in a documented way with computer actionable policy.

Metadata standards directory: Community curated standards catalogue for metadata interoperability

23 Things: Libraries For Research Data An overview of practical, free, online resources and tools that users can immediately leverage to incorporate research data management into librarianship.

Legal Interoperability of Research Data Principles and Implementation Guidelines: a set of principles and practical implementation guidelines offered as high-level guidance to all members of the research community —the funders, managers of data centers, librarians, archivists, publishers, policymakers, university administrators, individual researchers, and their legal counsel.

The Scholix initiative a high level interoperability framework for exchanging information about the links between scholarly literature and data. It aims to build an open information ecosystem to understand systematically what data underpins literature and what literature references data.

RDA Outputs

23 Things: Libraries For Research Data An overview of practical, free, online resources and tools that users can immediately leverage to incorporate research data management into librarianship.

Legal Interoperability of Research Data Principles and Implementation Guidelines: a set of principles and practical implementation guidelines offered as high-level guidance to all members of the research community —the funders, managers of data centers, librarians, archivists, publishers, policymakers, university administrators, individual researchers, and their legal counsel.

The Scholix initiative a high level interoperability framework for exchanging information about the links between scholarly literature and data. It aims to build an open information ecosystem to understand systematically what data underpins literature and what literature references data.

23 Things: Libraries for Research Data

An overview of practical, free, online resources and tools that you can begin using today to incorporate research data management into your practice of librarianship.

RESEARCH DATA ALLIANCE

Research Data Sharing Without Barriers

Learning Resources

Librarians are learning how to apply the principles of library science to solve problems and to provide new services related to research data.

1. A **"top ten" list of recommendations** for libraries to get started with research data management from LIBER, http://bit.ly/RDAthing1

2. Relevant concepts are presented and mapped in the **e-Science Thesaurus**, http://bit.ly/RDAthing2

3. Understanding the life of research data with the **DCC Curation Lifecycle Model**, http://bit.ly/RDAthing3

4. **MANTRA** online training modules for librarians,

Learning Resources
Data Reference and Outreach
Data Management Plans
Data Literacy
Citing Data
Data Licensing and Privacy
Digital Preservation
Data Repositories
and a Community of Practice

..to help librarians engage in research data management!

Data Reference & Outreach

Librarians are answering questions about data from patrons and conducting outreach to assess the data needs of their researchers and students.

7. Begin a conversation with a researcher about data by

---

10. Questions about data answered by experts on the **DataQ** forum, http://bit.ly/RDAthing10

Data Management Plans

Librarians are becoming familiar with funder requirements and consulting with researchers to help them write and implement effective data management plans.

11. One example is the **DMPTool** that lists funder requirements in the United States and builds a plan by asking the researcher to answer a series of questions. Other countries such as the U.K. and Canada have similar tools, http://bit.ly/RDAthing11

Data Literacy

Librarians are including data in their information literacy instruction, to

23 Things: Libraries for Research Data

RDA RESEARCH DATA ALLIANCE

An overview of practical, free, online resources and tools that you can begin using today to incorporate research data management into your practice of librarianship.

Research Data Sharing Without Barriers

Learning Resources

Librarians are learning how to apply the principles of library science to solve problems and to provide new services related to research data.

1. A "top ten" list of **recommendations** for libraries to get started with research data management from LIBER, http://bit.ly/RDAthing1

2. Relevant concepts are presented and mapped in the **e-Science Thesaurus**, http://bit.ly/RDAthing2

3. Understanding the life of research data with the **DCC Curation Lifecycle Model**, http://bit.ly/RDAthing3

4. **MANTRA** online training modules

Learning Resources
Data Reference and Outreach
Data Management Plans
Data Literacy
Citing Data
Data Licensing and Privacy
Digital Preservation
Data Repositories
and a Community of Practice

..to help librarians engage in research data management!

Data Reference & Outreach

Librarians are answering questions about data from patrons and conducting outreach to assess the data needs of their researchers and students.

10. Questions about data answered by experts on the **DataQ** forum, http://bit.ly/RDAthing10

Data Management Plans

Librarians are becoming familiar with funder requirements and consulting with researchers to help them write and implement effective data management plans.

11. One example is the **DMPTool** that lists funder requirements in the United States and builds a plan by asking the researcher to answer a series of questions. Other countries such as the U.K. and Canada have similar tools, http://bit.ly/RDAthing11

Data Literacy

Available in 11 languages

https://www.rd-alliance.org/group/libraries-research-data-ig/outcomes/23-things-libraries-research-data-supporting-output

RDA Outputs

23 Things: Libraries For Research Data An overview of practical, free, online resources and tools that users can immediately leverage to incorporate research data management into librarianship.

Legal Interoperability of Research Data Principles and Implementation Guidelines: a set of principles and practical implementation guidelines offered as high-level guidance to all members of the research community —the funders, managers of data centers, librarians, archivists, publishers, policymakers, university administrators, individual researchers, and their legal counsel.

The Scholix initiative a high level interoperability framework for exchanging information about the links between scholarly literature and data. It aims to build an open information ecosystem to understand systematically what data underpins literature and what literature references data.

# Introducing SCHOLIX: A framework for Scholarly Link Exchange

A new framework presenting a vision and guidelines for linking research data and literature using a common, global approach

- An evolving **lightweight set of Guidelines** to increase **interoperability** rather than a normative standard.

- The **consensus** achieved by the various stakeholder groups in the research data landscape – including data centers, publishers, Crossref, DataCite, OpenAIRE, and many others

SCHOLIX

See also http://www.scholix.org/about & http://bit.ly/29tdGNU
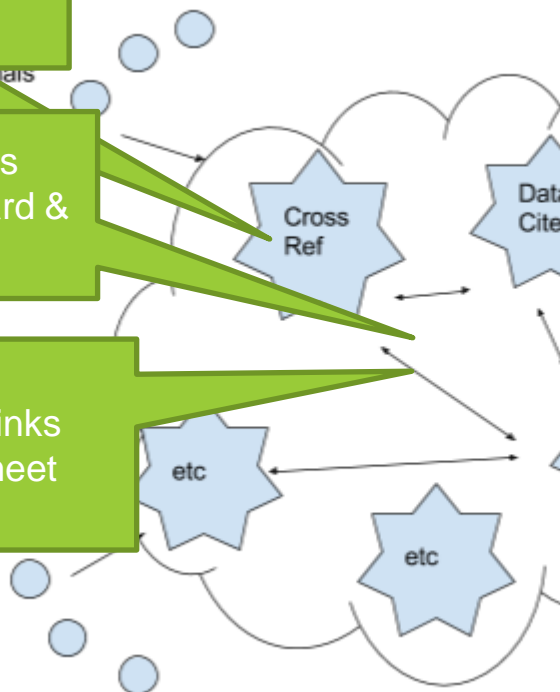
SCHOLIX: The Multi-Hub model
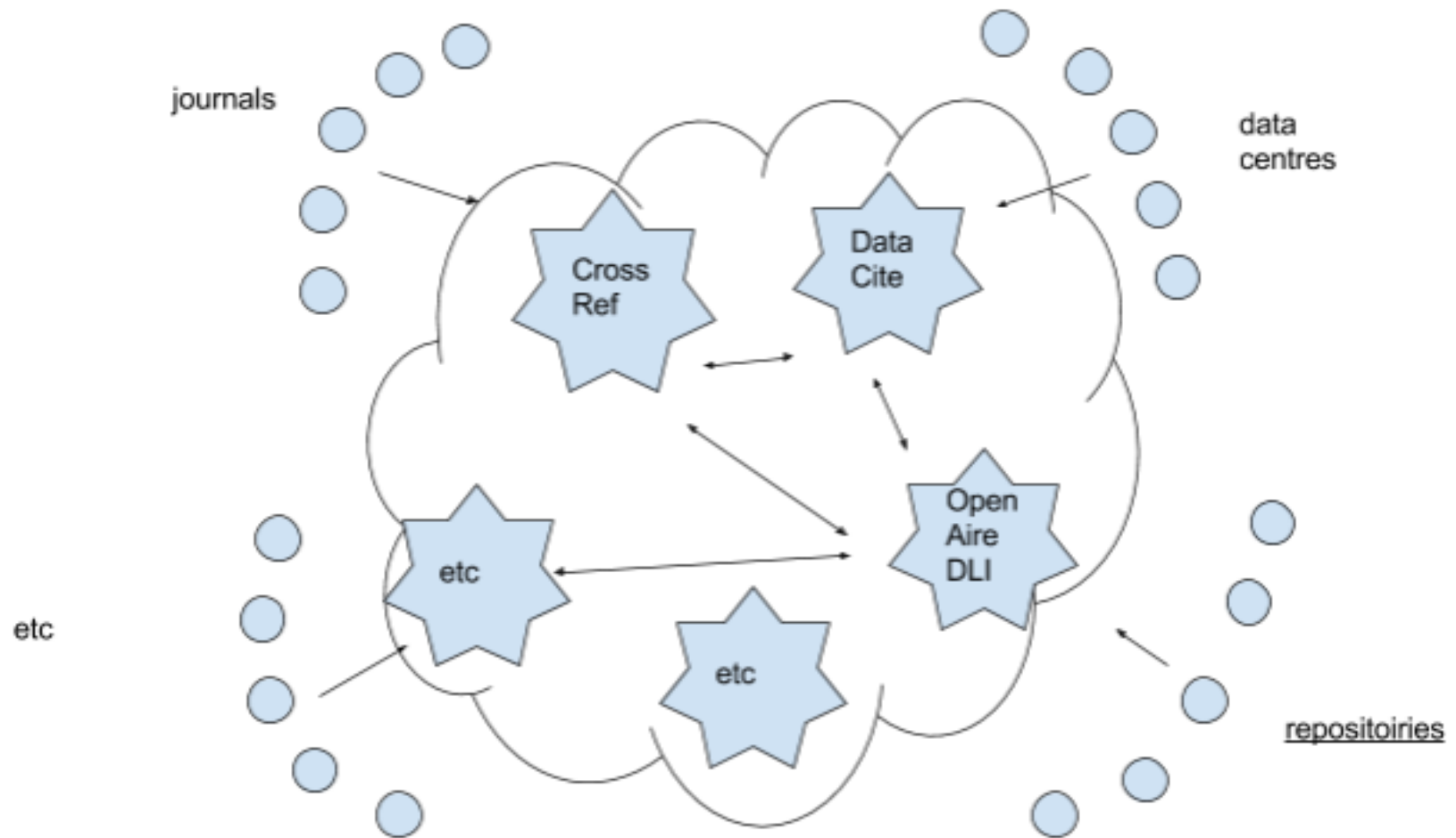
SCHOLIX: The Multi-Hub model

Link contributors announce link to a "hub" of their choice

Hubs convert/normalize links into a common data standard & share with other hubs

Services on top of this infrastructure aggregate links and provide services to meet specific use cases

Cross Ref

Data Cite

etc

etc

- Universal: cross-disciplinary, global

- Inclusive and participatory: supported by all stakeholder groups

- Open and non-discriminatory

- Quality through meticulous provenance and metadata (*not* "filtering at the gate")

- Standards-based

THE RESEARCH DATA ALLIANCE

www.rd-alliance.org

building the social and technical bridges that enable open sharing of data

15 FLAGSHIP OUTPUTS

of which 4 ICT Technical Specifications

75 ADOPTION CASES
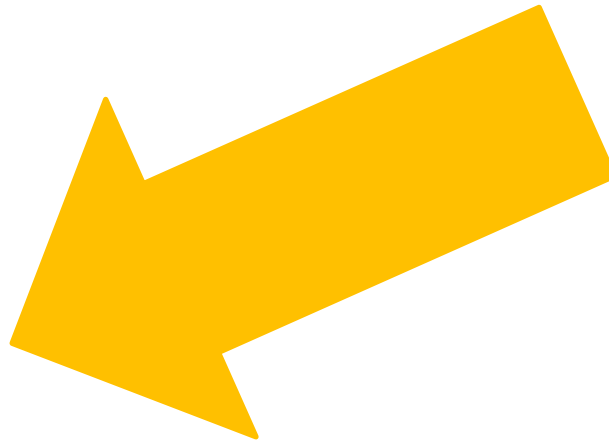
across multiple disciplines, organisations & countries

72 GROUPS WORKING ON GLOBAL DATA INTEROPERABILITY CHALLENGES

of which 27 WORKING GROUPS & 45 INTEREST GROUPS

4,527 INDIVIDUAL MEMBERS FROM 115 COUNTRIES

66% Academia & Research
16% Public Administration
11% Enterprise & Industry

46 ORGANISATIONAL MEMBERS

RDA
RESEARCH DATA ALLIANCE

How does RDA work?

RDA members **exchange knowledge, share discoveries, discuss barriers and potential solutions, define policies, harmonise standards, and tackle other infrastructure challenges to enhance and facilitate global data sharing** through:

Birds of a Feather (BOF)   Preliminary conversations to gauge interest and support for creating a group, usually short-term

Interest Groups (IG)   Ongoing, topical groups operating at a high level

Working Groups (WG)   12-18 month targeted efforts to address a specific, scoped problem, often spun off from Interest Groups

THE RESEARCH DATA ALLIANCE

www.rd-alliance.org

building the social and technical bridges that enable open sharing of data

15 FLAGSHIP OUTPUTS

of which 4 ICT Technical Specifications

75 ADOPTION CASES

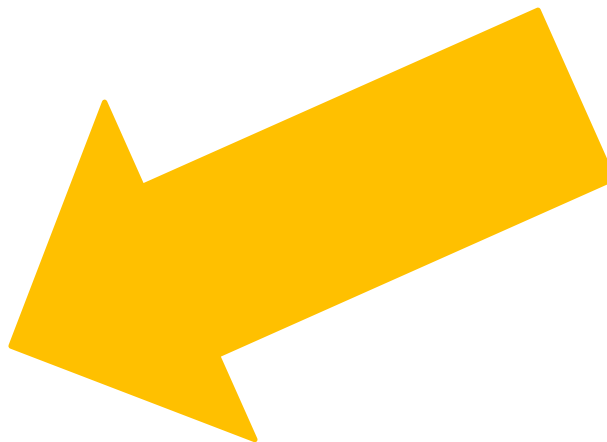across multiple disciplines, organisations & countries

72 GROUPS WORKING ON GLOBAL DATA INTEROPERABILITY CHALLENGES

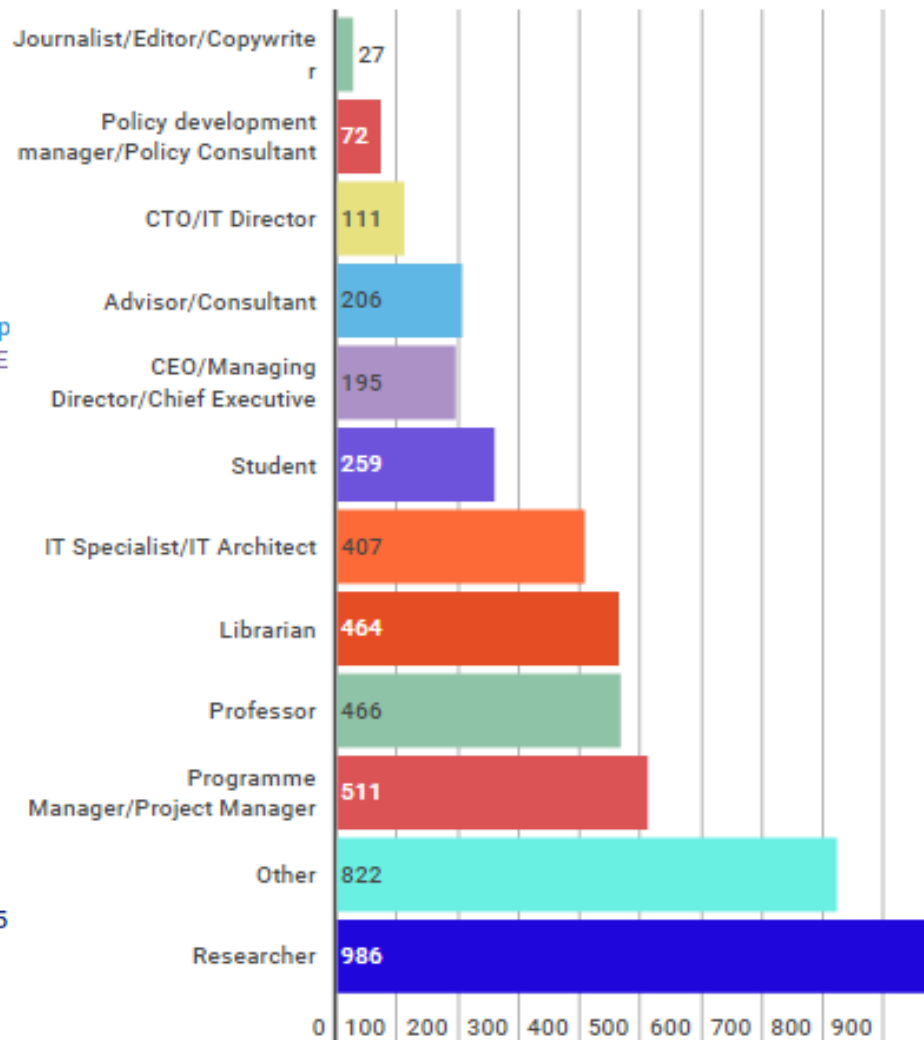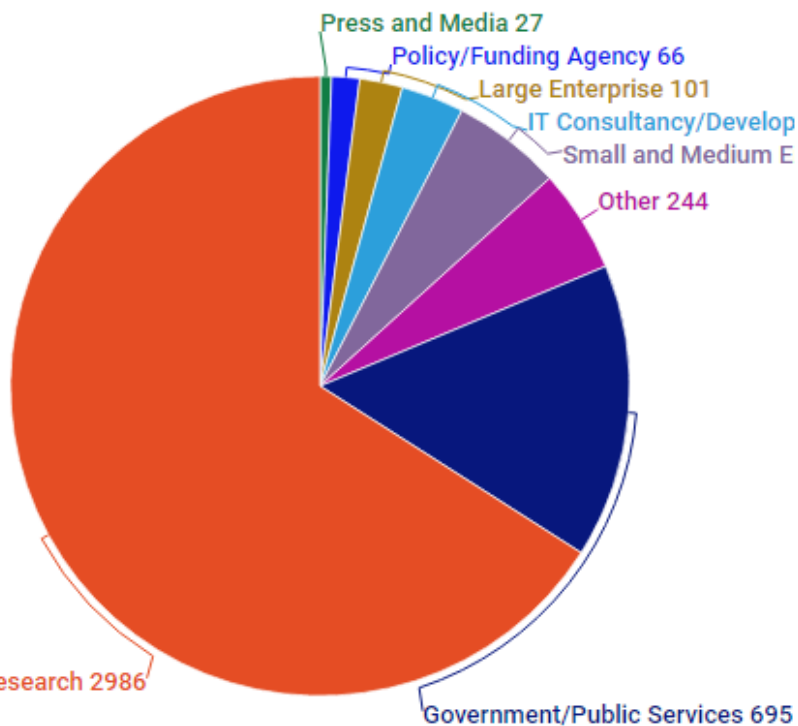*of which 27 WORKING GROUPS & 45 INTEREST GROUPS*

4,527 INDIVIDUAL MEMBERS FROM 115 COUNTRIES

66% Academia & Research

16% Public Administration

11% Enterprise & Industry

46 ORGANISATIONAL MEMBERS

RDA

RESEARCH DATA ALLIANCE

Who is RDA?

4527 total RDA community members from 115 countries

rd-alliance.org/about-rda   @resdatall

Who Can Join RDA?

Any individual or organization, regardless of profession or discipline, with an interest in reducing the barriers to data sharing and exchange and who agrees to RDA's guiding principles of:

- *Openness*
- *Consensus*
- *Balance*
- *Harmonization*
- *Community-driven*
- *Non-profit and technology-neutral*

Individual Membership is free @
http://www.rd-alliance.org/user/register

Leveraging RDA

PUSH: Domains, regions → RDA

◦ Bring domain and regional data efforts and issues from communities to the international, interdisciplinary stage

◦ Provide greater visibility and opportunities to community researchers and resources

PULL: RDA → Domains, regions

◦ Deploy RDA infrastructure in communities to accelerate the development of data sharing and data infrastructure

◦ Bring new ideas and collaborations to community researchers to enhance innovation and competitiveness
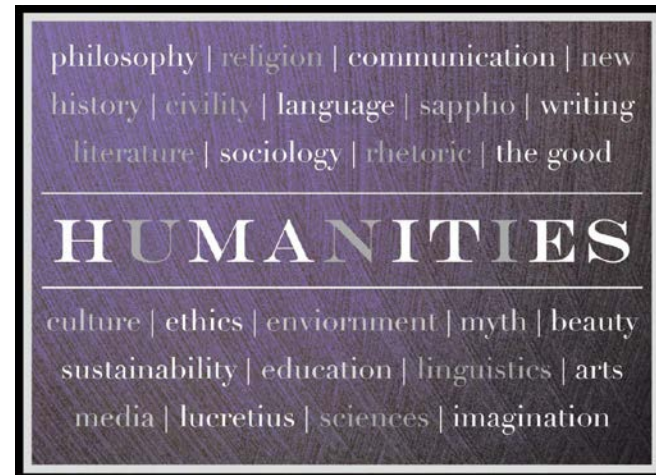
Potential Applications to Transportation

Potential Applications to Transportation

Potential Applications to Transportation

Figure derived from: KSBaker and F.Millerand, 2010. Infrastructuring Ecology: Challenges in Achieving Data Sharing. In Collaboration in the New Life Sciences. J.Parker, N.Vermeulen, and B.Penders (eds), Ashgate, Surrey, England: p. 111-138
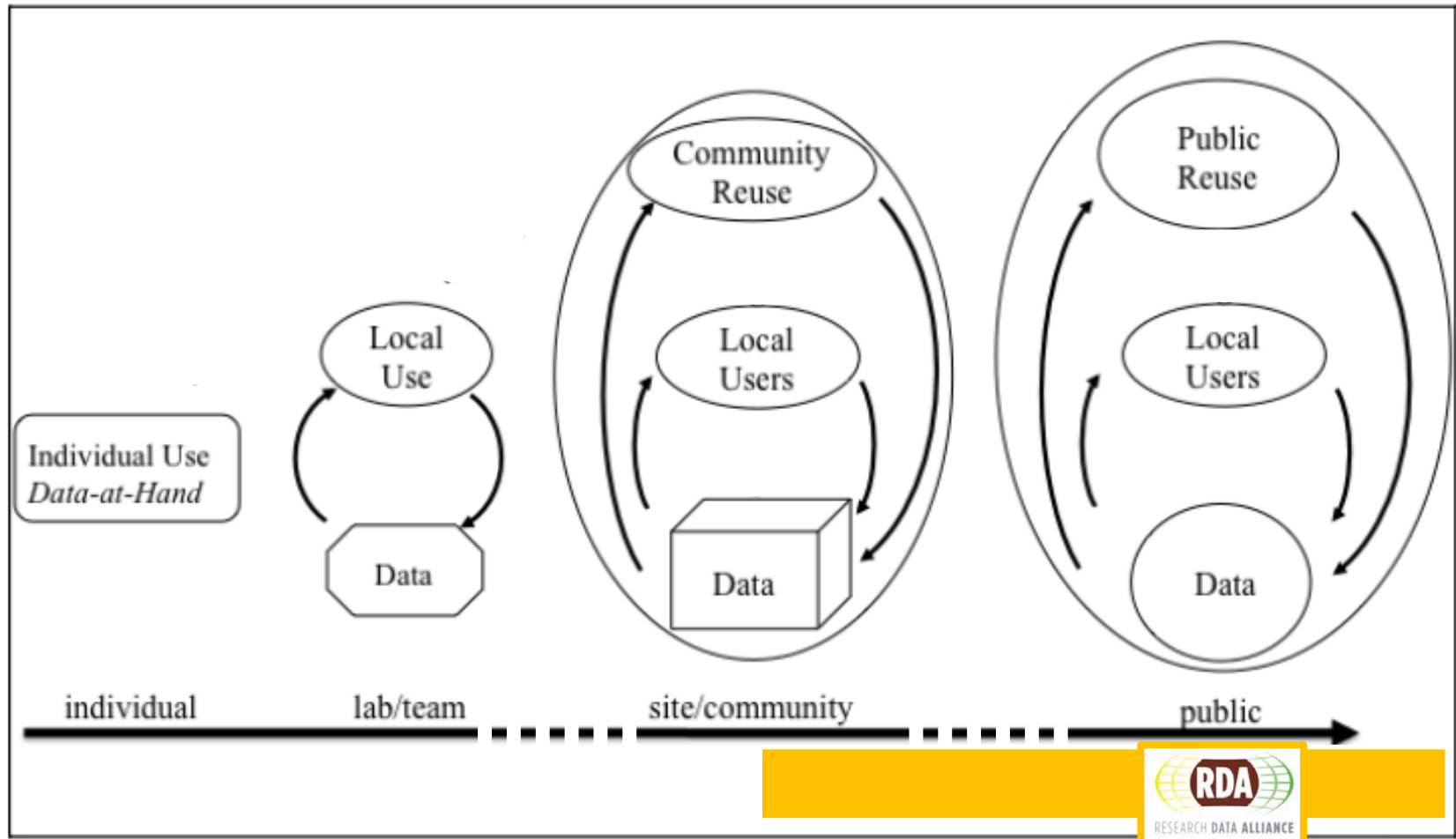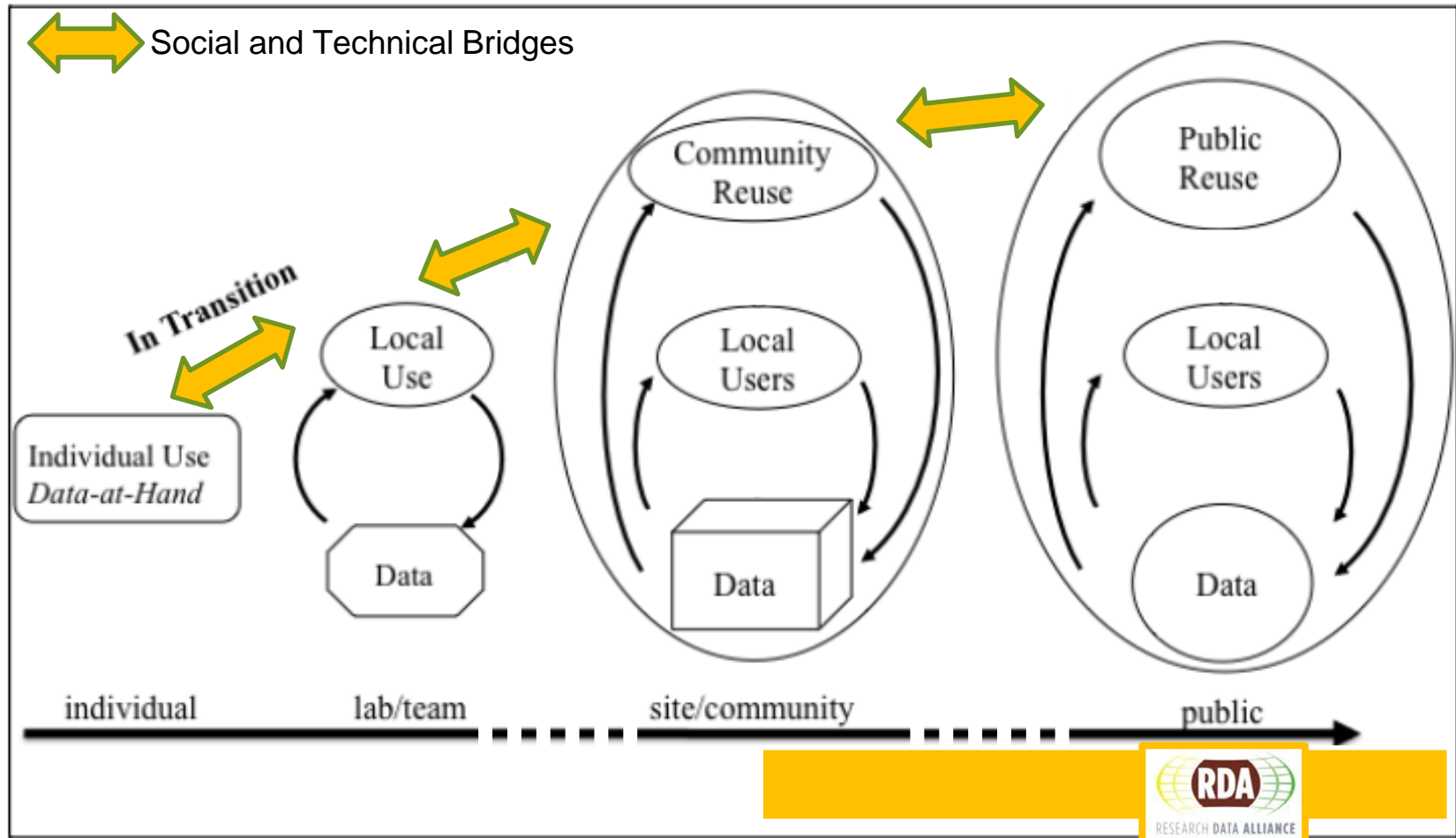
Potential Applications to Transportation

Consider adopting RDA Recommendations in your work:

PID Information Types: a common protocol for providers and users of persistent ID services worldwide.

Practical Policy: defining best practices of how to deal with data automatically and in a documented way with computer actionable policy.

◦**Metadata standards directory**: Community curated standards catalog for metadata interoperability

◦**Data Citation**: defining mechanisms to reliably cite dynamic data

RDA and Transportation

What RDA can bring to the table

- 4500+ people working to advance data topics
  - 20+ international, interdisciplinary Working Groups
  - Examples: repository interoperability, metadata standards, citation, brokering, persistent identifiers, privacy implications of data, preservation
- A neutral venue for resolving cross-boundary challenges
  - National, domain, community, agency, sector, etc.
- Administrative and support services for achieving consensus, adoption, and implementation

What the Transportation community can bring

- An community of diverse stakeholders
- A critical resource collection valuable to many other communities

Let's work together!

RESEARCH DATA ALLIANCE

RDA 9th Plenary Meeting
Data Infrastructures for Open Science
5-7 April 2017, Barcelo Sants Hotel,
Barcelona, Spain

Organised by Barcelona Supercomputing Center (BSC) with the support of RDA Europe

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

RDA
RESEARCH DATA ALLIANCE
EUROPE

https://www.rd-alliance.org/plenaries/rda-ninth-plenary-meeting-barcelona

RDA
RESEARCH DATA ALLIANCE

RDC ♦ DRC
Research Data Canada · Données de Recherche Canada

Université
de Montréal

Montréal, Canada
19-21 September 2017
19 au 21 septembre 2017

Plenary Meeting
Conférence plénière

https://www.rd-alliance.org/plenaries/rda-tenth-plenary-meeting-montreal-canada

Thank you!

Lynn Yarmey

Director of RDA/US Community Development

Yarmel@rpi.edu
@LynnYarmey

RDA US

Twitter - **@RDA_US**

RDA Global

Email - enquiries@rd-alliance.org

Web - www.rd-alliance.org

Twitter - @resdatall
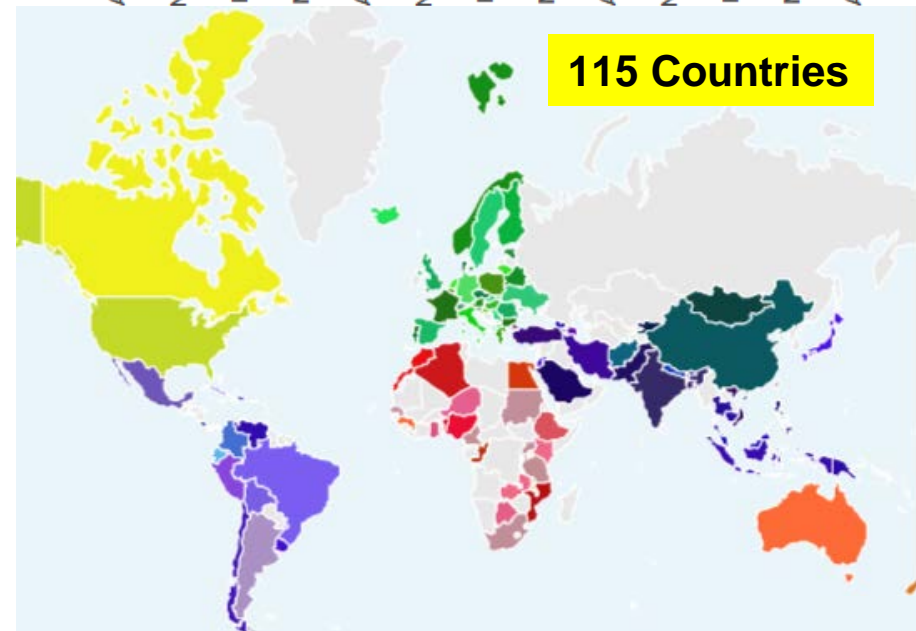
LinkedIn -
www.linkedin.com/in/ResearchDataAlliance

Slideshare -
http://www.slideshare.net/ResearchDataAlliance

Facebook -
https://www.facebook.com/pages/Research-Data-Alliance/459608890798924

RDA worldwide growth

Total RDA community members 4527

Bar chart values (left to right):
- Aug/Oct: 989
- Nov/Jan: 1272
- Feb/Apr: 1654
- May/Jul: 2046
- Aug/Oct: 2402
- Nov/Jan: 2634
- Feb/Apr: 2877
- May/Jul: 3122
- Aug/Oct: 3431
- Nov/Jan: 3694
- Feb/Apr: 4016
- May/Jul: 4273
- Aug/Oct: 4527

115 Countries

Pie chart:
- North America 33%
- South America 1%
- Africa 3%
- Asia 9%
- Australasia 5%
- Europe 48%

RESEARCH DATA ALLIANCE

Selected RDA Recommendations/outputs

Working Group	Outputs	Impact	Adopters
Dynamic Data Citation Working Group *(data consumer, social solution)*	Dynamic-data citation methodology that supports efficient processing of data and linking from publications	Researchers can reference precise subsets of changing data	NERC, ESIP, CLARIN, Virtual Atomic and Molecular Data Centre
Data Type Registries Working Group *(data provider, technical solution)*	Data type model and prototype registry	Provides machine-readable and researcher-accessible registries of data types that support the accurate use of data	CNRI, International DOI Foundation, Materials Genome Initiative, Deep Carbon Observatory, EUDAT
Wheat Data Interoperability Working Group *(data consumer, technical solution)*	Common framework for Wheat Data Terminology to enable interoperability between distinct data collections	Semantically linked terms describing wheat data so researchers can share harvest and related information between data sets and communities	Wheat Initiative Information System, FAO AIMS, INRA

RDA Recommendations & Outputs

THE RDA OUTCOMES LEGEND

Recommendations: are the flagship outputs of RDA. They are RDA's equivalent of the "specifications" or "standards" that other organisations create and endorse. The process for creating and endorsing these is already defined.

Supporting Outputs: are the outputs of RDA WGs and IGs that are fruit of RDA work, but are not necessarily adoptable bridges. "Upon request", these sort of outputs go through a community comment period and if no major objections or gaps are identified they get the RDA Brand.

Other Outputs: include workshop reports, published articles, survey results, etc. Anything a WG or IG wants to register and report. Upon request, these are published and discoverable on the RDA website but have no level of endorsement.

RDA Interest (IG) & Working Groups (WG)

Total 72 groups:
27 Working Groups &
45 Interest Groups

Domain Research

Agrisemantics WG

BioSharing Registry: connecting data policies, standards & databases in life sciences WG

Fisheries Data Interoperability WG

On-Farm Data Sharing (OFDS) Working Group

Rice Data Interoperability WG

Wheat Data Interoperability WG

Agriculture Data IG (IGAD)

Biodiversity Data Integration IG

Chemistry Research Data IG

Digital Practices in History and Ethnography IG

Geospatial IG

Global Water Information IG

Health Data IG

Libraries for Research Data IG

Marine Data Harmonization IG

Metabolomics Data Interoperability IG

Quality of Urban Life IG

RDA/CODATA Materials Data, Infrastructure & Interoperability

Research data needs of the Photon and Neutron Science community IG

Structural Biology IG

Community Needs

RDA/CODATA Summer Schools in Data Science and Cloud Computing in the Developing World WG

Teaching TDM on Education and Skill Development WG

Archives & Records Professionals for Research Data IG

Data for Development IG

Development of Cloud Computing Capacity and Education in Developing World Research IG

Education and Training on handling of research data IG

Engagement IG

Ethics and Social Aspects of Data IG

RDA Interest (IG) & Working Groups (WG)

Total 72 groups:
27 Working Groups &
45 Interest Groups

Reference and Sharing

Data Citation WG

Data Description Registry Interoperability WG

Data Security and Trust WG

Empirical Humanities Metadata WG

RDA / WDS Publishing Data Bibliometrics WG

Research Data Collections WG

QoS-DataLC Definitions WG

International Materials Resource Registries WG

National Data Services IG

RDA/CODATA Legal Interoperability IG

Reproducibility IG

New Paradigms for Data Discovery IG

Partnership Groups

RDA / TDWG Metadata Standards for attribution of physical and digital collections stewardship WG

RDA/NISO Privacy Implications of Research Data Sets WG

RDA/WDS Scholarly Link Exchange WG

Repository Audit and Certification DSA–WDS Partnership WG

RDA/WDS Publishing Data IG

ELIXIR Bridging Force IG

RDA Interest (IG) & Working Groups (WG)

Total 72 groups:
27 Working Groups &
45 Interest Groups

Data Stewardship and Services

Brokering Framework WG

Brokering Governance WG

RDA / WDS Publishing Data Services WG

RDA / WDS Publishing Data Workflows WG

Active Data Management Plans IG

Data in Context IG

Data Rescue IG

Domain Repositories IG

Long tail of research data IG

Preservation e-Infrastructure IG

RDA/WDS Certification of Digital Repositories IG

RDA/WDS Publishing Data Cost Recovery for Data Centres IG

Repository Platforms for Research Data IG

Research Data Provenance IG

Virtual Research Environments IG

Foundational Infrastructure

Array Database WG
Data Foundation and Terminology WG
Data Type Registries WG
Metadata Standards Catalog WG
Metadata Standards Directory WG
PID Information Types WG
Practical Policy WG
Data Fabric IG

Data Foundations and Terminology IG
Big Data IG
Brokering IG
Federated Identity Management IG
Metadata IG
PID IG
Service Management IG
Vocabulary Services IG

Get your work done

Examples of how other communities have leveraged RDA:

Wheat Data Interoperability Working Group - impacting the discoverability, reusability and interoperability of wheat data by building a common framework for describing, representing linking and publishing wheat data

Chemistry Data Interest Group - foster diverse professional exchange on issues particular to data originating from the field of chemistry

Health Data Interest Group - This IG deals mainly with the vertical of Health Data, while other groups dealing with privacy, security and trust are horizontal with potential use cases from several areas.